

CLAIMS:

1. A speech synthesis apparatus comprising:
 - means for inputting of natural speech,
 - means for processing the natural speech to provide personalized speech segments,
 - 5 - means for synthesizing of speech based on the personalized speech segments.
2. The speech synthesis apparatus of claim 1, the means for processing the natural speech comprising means for extracting of speech segments from natural speech.
- 10 3. The speech synthesis apparatus of claims 1 or 2 further comprising:
 - a speech segments database for storing of speech segments, the speech segments having marker information assigned thereto,
 - means for finding a best match of a speech segment in the speech segments database and natural speech,
 - 15 - means for copying the marker information after the best match has been performed to the natural speech.
4. The speech synthesis apparatus of claim 3, the means for finding a best match being adapted to perform a dynamic time warping (DTW) type method.
- 20 5. The speech synthesis apparatus of any one of the preceding claims 1 to 4 further comprising a personalized speech segments database (114) for storing of extracted speech segments, the extracted speech segments having marker information assigned thereto.
- 25 6. The speech synthesis apparatus of any of the preceding claims 1 to 5 further comprising means for storing a list of words to be spoken by a speaker to provide the personalized speech segments.

7. The speech synthesis apparatus of any one of the preceding claims 1 to 6 further comprising a user interface for display of words to be spoken by a user.

8. The speech synthesis apparatus of any one of the preceding claims 1 to 7 further comprising means for rendering of words to be spoken prior to inputting of the natural speech.

9. The speech synthesis apparatus of any one of the preceding claims 1 to 8 further comprising:

- a set of personalized speech segments databases for different speakers,
- means for selecting one of the personalized speech segments databases from the set of personalized speech segments databases.

10. The speech synthesis apparatus of any one of the preceding claims 1 to 9 further comprising means for exporting of the personalized speech segments.

11. The speech synthesis apparatus of any one of the preceding claims 1 to 10, the natural speech to be inputted comprising a list of nonsense words.

12. The speech synthesis apparatus of any one of the preceding claims 1 to 11, the speech segments being diphones, triphones and/or polyphones.

13. The speech synthesis apparatus of any one of the preceding claims 1 to 12, the means for synthesizing of speech being adapted to perform the speech synthesis by means of a PSOLA type method.

14. The speech synthesis apparatus of any one of the preceding claims 1 to 13, further comprising control means for providing text to the means for synthesizing of speech.

15. A consumer device, such as an audio, video, household, camera, computer, telecommunication, car navigation and/or personal digital assistant device, comprising a speech synthesis apparatus in accordance with any one of the preceding claims 1 to 14 for providing of a personalized natural speech output.

16. A method of speech synthesis comprising the steps of:

- inputting of natural speech into a consumer device,
- processing of the natural speech by the consumer device to provide

personalized speech segments,

- 5 - synthesizing of text-to-speech to provide a personalized speech output based on the personalized speech segments for text to be outputted by the consumer device.

17. The method of claim 16 further comprising extracting of speech segments from the natural speech.

10

18. The method of claims 16 or 17 further comprising the steps of:

- identifying a best matching speech segment for inputted natural speech in a database, the database comprising speech segments having marker information assigned thereto,

- 15 - assigning the marker information of the identified best matching speech segment to the natural speech.

19. The method of claims 16, 17 or 18, whereby a dynamic time warping (DTW) type method is employed for identification of the best matching speech segment.

20

20. A computer program product, such as a digital storage medium, comprising computer program means for performing the steps of:

- inputting of natural speech into a consumer device,
- processing of the natural speech within the consumer device to provide

25 personalized speech segments,

- synthesizing of text-to-speech to provide a personalized speech output based on the personalized speech segments for text to be outputted by the consumer device.